```
                        GenCore version 5.1.6
               Copyright (c) 1993 - 2004  Compugen Ltd.


OM protein - protein search, using sw model

Run on:          April 15, 2004, 13:37:36 ; Search time 50.2 Seconds
                                           (without alignments)
                                           1412.738 Million cell updates/sec

Title:           US-08-737-319(1)   in PROTEIN DATABASES
Perfect score:   1341
Sequence:        1 MSMPNIVPPAEVRTEGLSLE..........KGEVDAKSLEDLSDNKVWKM 251

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters:      1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                          PROTEIN SEQ 1
Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_29Jan04:*
                 1:  geneseqp1980s:*
                 2:  geneseqp1990s:*
                 3:  geneseqp2000s:*
                 4:  geneseqp2001s:*
                 5:  geneseqp2002s:*
                 6:  geneseqp2003as:*
                 7:  geneseqp2003bs:*
                 8:  geneseqp2004s:*

       Pred. No. is the number of results predicted by chance to have a
       score greater than or equal to the score of the result being printed,
       and is derived by analysis of the total score distribution.


                              SUMMARIES
               %
Result        Query                               Appl.      earliest
  No.   Score Match Length DB  ID                     Description  12/22/95
----------------------------------------------------------------------------
      1  1341 100.0   251   2  AAW00278 WO 96/28545   Aaw00278 Peptide u
      2  1341 100.0   251   2  AAW22501 WO 97/23633   Aaw22501 Phaffia d
      3   819  61.1   281   7  ADB70021               Adb70021 C. neofor
      4 640.5  47.8   271   6  ABJ25766               Abj25766 Aspergill
      5   629  46.9   284   5  ABP73637               Abp73637 Candida a
      6 627.5  46.8   271   6  ABJ26366               Abj26366 Aspergill
      7 579.5  43.2   228   5  ABP62828               Abp62828 Human pol
      8 579.5  43.2   228   6  ABU89730               Abu89730 Protein d
      9 579.5  43.2   228   7  ADE54597               Ade54597 Human Pro
```

10	578.5	43.1	288	2	AAW00280		Aaw00280	Peptide u
11	575.5	42.9	288	5	ABG93897		Abg93897	Saccharom
12	560	41.8	227	7	ADE54595		Ade54595	Rat Prote
13	547	40.8	234	3	AAY54297		Aay54297	Amino aci
14	545.5	40.7	295	3	AAY54296		Aay54296	Amino aci
15	541	40.3	238	3	AAY54298		Aay54298	Amino aci
16	539	40.2	238	5	ABG93909		Abg93909	Oryza sat
17	536	40.0	287	5	ABG93895		Abg93895	Clarkia b
18	535.5	39.9	295	5	ABG93907		Abg93907	Adonis pa
19	535	39.9	234	5	ABG93908		Abg93908	Adonis pa
20	532.5	39.7	284	2	AAW35349		Aaw35349	Arabidops
21	532.5	39.7	284	3	AAY54285		Aay54285	An Arabid
22	530.5	39.6	284	3	AAY77972		Aay77972	A. thalia
23	528.5	39.4	225	3	AAG42236		Aag42236	Arabidops
24	528.5	39.4	233	3	AAG42235		Aag42235	Arabidops
25	528.5	39.4	291	3	AAG42234		Aag42234	Arabidops
26	527.5	39.3	280	5	ABG93905		Abg93905	Lactuca s
27	526.5	39.3	229	3	AAY54295		Aay54295	Amino aci
28	526.5	39.3	229	5	ABG93906		Abg93906	Lactuca s
29	525.5	39.2	284	5	ABG93894		Abg93894	Arabipsos
30	523.5	39.0	233	3	AAY54299		Aay54299	Amino aci
31	523.5	39.0	261	2	AAW35348		Aaw35348	Arabidops
32	523.5	39.0	261	3	AAY54286		Aay54286	An Arabid
33	520.5	38.8	261	5	ABG93896		Abg93896	Arabipsos
34	520.5	38.8	280	3	AAY54294		Aay54294	Amino aci
35	517.5	38.6	232	3	AAY54293		Aay54293	Amino aci
36	515.5	38.4	232	5	ABG93904		Abg93904	Tagetes e
37	511.5	38.1	234	4	AAB30549		Aab30549	An isopen
38	511.5	38.1	234	4	AAB99771		Aab99771	Hevea bra
39	510.5	38.1	233	5	ABG93910		Abg93910	Arabipsos
40	510	38.0	232	3	AAY90229		Aay90229	Marigold
41	500	37.3	307	3	AAY54302		Aay54302	Amino aci
42	499	37.2	307	5	ABG93913		Abg93913	Chlamydom
43	494.5	36.9	304	3	AAY54301		Aay54301	Amino aci
44	494	36.8	259	2	AAW00279		Aaw00279	Peptide u
45	493	36.8	305	2	AAW35350		Aaw35350	Haematoco

ALIGNMENTS

RESULT 1
AAW00278
ID   AAW00278 standard; protein; 251 AA.
XX
AC   AAW00278;
XX
DT   16-OCT-2003 (revised)
DT   14-JAN-1997 (first entry)
XX
DE   Peptide used to increase carotenoid yield.
XX
KW   Carotenoid; beta-carotene; phytoene; lycopene; zeaxanthin; astaxanthin;
KW   Phaffia rhodozyma; Haematococcus pluvialis; Saccharomyces cerevisiae.
XX
OS   Xanthophyllomyces dendrorhous.
XX

```
                         GenCore version 5.1.6
                 Copyright (c) 1993 - 2004  Compugen Ltd.


OM protein - protein search, using sw model

Run on:          April 15, 2004, 13:50:32 ; Search time 13.7804 Seconds
                                           (without alignments)
                                           1752.060 Million cell updates/sec

Title:           US-08-737-319-1
Perfect score:   1341
Sequence:        1 MSMPNIVPPAEVRTEGLSLE..........KGEVDAKSLEDLSDNKVWKM 251


Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5


Searched:        283366 seqs, 96191526 residues


Total number of hits satisfying chosen parameters:       283366


Minimum DB seq length: 0
Maximum DB seq length: 2000000000


Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries


Database :       PIR_78:*
                 1:  pir1:*
                 2:  pir2:*
                 3:  pir3:*
                 4:  pir4:*


     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                     %
Result             Query
  No.     Score    Match Length DB   ID                    Description
--------------------------------------------------------------------------------
      1      616    45.9    227   2  A56442                isopentenyl-diphos
      2    579.5    43.2    228   2  A53028                isopentenyl-diphos
      3    578.5    43.1    288   2  A34440                isopentenyl-diphos
      4      536    40.0    290   2  S49588                isopentenyl-diphos
      5    530.5    39.6    233   2  S71369                isopentenyl-diphos
      6    530.5    39.6    284   2  S71370                isopentenyl-diphos
      7      500    37.3    307   2  T07979                isopentenyl-diphos
      8      498    37.1    293   2  T51248                isopentenyl-diphos
      9      494    36.8    293   2  T46812                isopentenyl-diphos
     10      493    36.8    305   2  T52027                isopentenyl-diphos
     11    403.5    30.1    831   2  S44843                K06H7.3 protein -
     12    230.5    17.2    219   2  T03399                isopentenyl-diphos
     13      209    15.6    197   2  T35275                probable IPP isome
```

```
                        GenCore version 5.1.6
                 Copyright (c) 1993 - 2004  Compugen Ltd.


OM protein - protein search, using sw model

Run on:          April 15, 2004, 13:46:31 ; Search time 9.84314 Seconds
                                            (without alignments)
                                            1327.789 Million cell updates/sec


Title:           US-08-737-319-1
Perfect score:   1341
Sequence:        1 MSMPNIVPPAEVRTEGLSLE..........KGEVDAKSLEDLSDNKVWKM 251


Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5


Searched:        141681 seqs, 52070155 residues


Total number of hits satisfying chosen parameters:      141681


Minimum DB seq length: 0
Maximum DB seq length: 2000000000


Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries


Database :       SwissProt_42:*


     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.


                            SUMMARIES
                      %
Result           Query
   No.    Score  Match Length DB   ID                          Description
   ------------------------------------------------------------------------------
      1    1337   99.7    251  1    IDI1_PHARH  1997  GenBank O42641 phaffia rho
      2     616   45.9    227  1    IDI1_SCHPO              Q10132 schizosacch
      3   579.5   43.2    227  1    IDI1_HUMAN              Q13907 homo sapien
      4   578.5   43.1    287  1    IDI1_YEAST              P15496 saccharomyc
      5   567.5   42.3    227  1    IDI1_MESAU              O35586 mesocricetu
      6   562.5   41.9    227  1    IDI1_MOUSE              P58044 mus musculu
      7     560   41.8    227  1    IDI1_RAT                O35760 rattus norv
      8     536   40.0    287  1    IDI1_CLABR              Q39472 clarkia bre
      9   532.5   39.7    235  1    IDI1_CAMAC              O48964 camptotheca
     10   530.5   39.6    284  1    IDI2_ARATH              Q42553 arabidopsis
     11   528.5   39.4    233  1    IDI1_ARATH              Q38929 arabidopsis
     12     520   38.8    309  1    IDI2_CAMAC              O48965 camptotheca
     13   507.5   37.8    286  1    IDI2_CLABR              Q39471 clarkia bre
     14   506.5   37.8    290  1    IDI2_CLAXA              Q39664 clarkia xan
     15     442   33.0    227  1    IDI2_HUMAN              Q9bxs1 homo sapien
     16     209   15.6    197  1    IDI_STRCO               Q9x7q6 streptomyce
     17     205   15.3    197  1    IDI_STRAW               Q82mj7 streptomyce
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2004, 13:47:16 ; Search time 35.4353 Seconds
                                          (without alignments)
                                          2234.921 Million cell updates/sec

Title: US-08-737-319-1
Perfect score: 1341
Sequence: 1 MSMPNIVPPAEVRTEGLSLE..........KGEVDAKSLEDLSDNKVWKM 251

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   SPTREMBL_25:*
                1:   sp_archea:*
                2:   sp_bacteria:*
                3:   sp_fungi:*
                4:   sp_human:*
                5:   sp_invertebrate:*
                6:   sp_mammal:*
                7:   sp_mhc:*
                8:   sp_organelle:*
                9:   sp_phage:*
                10: sp_plant:*
                11: sp_rodent:*
                12: sp_virus:*
                13: sp_vertebrate:*
                14: sp_unclassified:*
                15: sp_rvirus:*
                16: sp_bacteriap:*
                17: sp_archeap:*

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	1341	100.0	251	3	Q9UUQ1	*1998*	Q9uuq1 phaffia rho
2	637	47.5	268	3	Q874I0		Q874i0 emericella
3	636	47.4	234	3	Q8NJL9		Q8njl9 aspergillus
4	579.5	43.2	228	4	Q86U81		Q86u81 homo sapien
5	547.5	40.8	295	10	Q9M6K9		Q9m6k9 adonis pala
6	547	40.8	267	10	Q9M6K8		Q9m6k8 adonis pala
7	541	40.3	238	10	Q8GVZ0		Q8gvz0 oryza sativ
8	541	40.3	254	10	Q9M6K4		Q9m6k4 oryza sativ
9	540	40.3	290	10	Q9AVG8		Q9avg8 nicotiana t
10	532.5	39.7	284	10	Q9M6K3		Q9m6k3 arabidopsis
11	529.5	39.5	235	10	Q9AVG7		Q9avg7 nicotiana t
12	529.5	39.5	263	10	Q9M6K6		Q9m6k6 lactuca sat
13	524	39.1	281	10	Q9M592		Q9m592 brassica ol
14	523.5	39.0	261	10	Q9M6K2		Q9m6k2 arabidopsis
15	522.5	39.0	226	10	Q8S3L7		Q8s3l7 melaleuca a
16	520.5	38.8	280	10	Q9M6K7		Q9m6k7 lactuca sat
17	518.5	38.7	245	10	Q9M6K5		Q9m6k5 tagetes ere
18	518	38.6	224	10	Q9FXR6		Q9fxr6 eucommia ul
19	518	38.6	239	5	Q9NH02		Q9nh02 dictyosteli
20	511.5	38.1	234	10	Q9S7C4		Q9s7c4 hevea brasi
21	510	38.0	262	10	Q9FV48		Q9fv48 tagetes ere
22	507.5	37.8	235	10	Q8S3L8		Q8s3l8 melaleuca a
23	507.5	37.8	235	10	Q84RZ8		Q84rz8 periploca s
24	500	37.3	307	10	O81691		O81691 chlamydomon
25	498	37.1	293	10	O81660		O81660 haematococc
26	494	36.8	293	10	Q9SAY0		Q9say0 haematococc
27	493	36.8	305	10	O81659		O81659 haematococc
28	489	36.5	256	5	Q9VDC2		Q9vdc2 drosophila
29	488	36.4	256	5	Q8MRH1		Q8mrh1 drosophila
30	440	32.8	227	11	Q8BFZ6		Q8bfz6 mus musculu
31	403.5	30.1	236	5	Q9BI81		Q9bi81 caenorhabdi
32	321	23.9	120	10	Q9AVI8		Q9avi8 youngia jap
33	320	23.9	120	10	Q94IE1		Q94ie1 euphorbia t
34	318	23.7	120	10	Q9AVG5		Q9avg5 sapium sebi
35	313	23.3	120	10	Q9AVJ5		Q9avj5 sonchus ole
36	304	22.7	120	10	Q9AYS6		Q9ays6 sonchus ole
37	301.5	22.5	121	10	Q9AVG6		Q9avg6 taraxacum j
38	301	22.4	119	10	Q9AVJ1		Q9avj1 metaplexis
39	266	19.8	229	5	Q8SSI5		Q8ssi5 encephalito
40	261	19.5	91	10	Q9LLB6		Q9llb6 daucus caro
41	249.5	18.6	197	2	Q8RTW6		Q8rtw6 uncultured
42	230.5	17.2	219	10	P93355		P93355 nicotiana t
43	221.5	16.5	503	16	Q83DT5		Q83dt5 coxiella bu
44	205	15.3	192	2	Q8KZ66		Q8kz66 uncultured
45	181	13.5	172	2	Q8KZ12		Q8kz12 uncultured

ALIGNMENTS

RESULT 1
Q9UUQ1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 15, .2004, 13:49:12 ; Search time 16.2412 Seconds
                                         (without alignments)
                                         797.856 Million cell updates/sec

Title: US-08-737-319-1
Perfect score: 1341
Sequence: 1 MSMPNIVPPAEVRTEGLSLE..........KGEVDAKSLEDLSDNKVWKM 251

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_AA:*
               1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
               2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
               3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
               4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
               5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
               6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                              SUMMARIES
                    %   USPN 6329141
Result   Query
   No.   Score Match Length DB ID   Description
----
     1   1341 100.0   251 4   US-09-091-725-21   Sequence 21, Appl
     2   578.5   43.1   288 1   US-08-624-125-19   Sequence 19, Appl
     3   578.5   43.1   288 4   US-08-937-155-19   Sequence 19, Appl
     4   578.5   43.1   288 4   US-09-323-998E-19   Sequence 19, Appl
     5   547   40.8   234 4   US-09-323-998E-40   Sequence 40, Appl
     6   545.5   40.7   295 4   US-09-323-998E-39   Sequence 39, Appl
     7   541   40.3   238 4   US-09-323-998E-41   Sequence 41, Appl
     8   536   40.0   287 1   US-08-624-125-17   Sequence 17, Appl
     9   536   40.0   287 4   US-08-937-155-17   Sequence 17, Appl
    10   536   40.0   287 4   US-09-323-998E-17   Sequence 17, Appl
    11   532.5   39.7   284 1   US-08-624-125-16   Sequence 16, Appl

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2004, 13:54:22 ; Search time 37.8961 Seconds
                                           (without alignments)
                                           1825.887 Million cell updates/sec

Title: US-08-737-319-1
Perfect score: 1341
Sequence: 1 MSMPNIVPPAEVRTEGLSLE..........KGEVDAKSLEDLSDNKVWKM 251

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published_Applications_AA:*
                 1:   /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
                 2:   /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
                 3:   /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
                 4:   /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
                 5:   /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
                 6:   /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
                 7:   /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
                 8:   /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
                 9:   /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
                 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
                 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
                 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
                 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
                 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
                 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
                 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
                 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
                 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

```
                                SUMMARIES
                       %
Result               Query
   No.    Score    Match  Length  DB   ID                              Description
       ---------------------------------------------------------------------------------
        1      819    61.1      281   15   US-10-320-797-3065         Sequence 3065, Ap
        2    640.5    47.8      271   14   US-10-128-714-3424         Sequence 3424, Ap
        3      629    46.9      284   14   US-10-032-585-7474         Sequence 7474, Ap
        4    627.5    46.8      271   14   US-10-128-714-8424         Sequence 8424, Ap
        5    579.5    43.2      228   12   US-10-363-616-265          Sequence 265, App
        6    579.5    43.2      273   15   US-10-369-493-12865        Sequence 12865, A
        7    578.5    43.1      288    9   US-09-323-998D-19        Sequence 19, Appl
        8    578.5    43.1      288   15   US-10-369-493-22482        Sequence 22482, A
        9      574    42.8      267   15   US-10-369-493-2346         Sequence 2346, Ap
       10      574    42.8      267   15   US-10-369-493-2351         Sequence 2351, Ap
       11      557    41.5      257   15   US-10-369-493-3496         Sequence 3496, Ap
       12    547.5    40.8      295   16   US-10-389-566-2294         Sequence 2294, Ap
       13      547    40.8      234    9   US-09-323-998D-40        Sequence 40, Appl
       14      547    40.8      267   16   US-10-389-566-2295         Sequence 2295, Ap
       15    545.5    40.7      295    9   US-09-323-998D-39        Sequence 39, Appl
       16      541    40.3      238    9   US-09-323-998D-41        Sequence 41, Appl
       17      541    40.3      238   16   US-10-389-566-1985         Sequence 1985, Ap
       18      536    40.0      287    9   US-09-323-998D-17        Sequence 17, Appl
       19    532.5    39.7      284    9   US-09-323-998D-16        Sequence 16, Appl
       20    530.5    39.6      284   14   US-10-342-224-96           Sequence 96, Appl
       21    529.5    39.5      263   16   US-10-389-566-2297         Sequence 2297, Ap
       22    526.5    39.3      229    9   US-09-323-998D-38        Sequence 38, Appl
       23      524    39.1      281   16   US-10-389-566-2306         Sequence 2306, Ap
       24    523.5    39.0      233    9   US-09-323-998D-42        Sequence 42, Appl
       25    523.5    39.0      261    9   US-09-323-998D-18        Sequence 18, Appl
       26    522.5    39.0      226   16   US-10-389-566-1492         Sequence 1492, Ap
       27      521    38.9      309   16   US-10-389-566-792          Sequence 792, App
       28    520.5    38.8      280    9   US-09-323-998D-37        Sequence 37, Appl
       29    520.5    38.8      280   16   US-10-389-566-2296         Sequence 2296, Ap
       30    518.5    38.7      245   16   US-10-389-566-2298         Sequence 2298, Ap
       31    517.5    38.6      232    9   US-09-323-998D-36        Sequence 36, Appl
       32    511.5    38.1      234   16   US-10-389-566-2210         Sequence 2210, Ap
       33      510    38.0      262   16   US-10-389-566-2458         Sequence 2458, Ap
       34    507.5    37.8      235   16   US-10-389-566-1491         Sequence 1491, Ap
       35    506.5    37.8      290   16   US-10-389-566-2247         Sequence 2247, Ap
       36      500    37.3      307    9   US-09-323-998D-45        Sequence 45, Appl
       37      498    37.1      293   16   US-10-389-566-2102         Sequence 2102, Ap
       38      494    36.8      293   16   US-10-389-566-2125         Sequence 2125, Ap
       39      493    36.8      305    9   US-09-323-998D-14        Sequence 14, Appl
       40      493    36.8      305    9   US-09-323-998D-44        Sequence 44, Appl
       41      493    36.8      305   16   US-10-389-566-2101         Sequence 2101, Ap
       42      490    36.5      194    9   US-09-925-302-585        Sequence 585, App
       43      490    36.5      194   12   US-09-925-302-585          Sequence 585, App
       44      490    36.5      293    9   US-09-323-998D-15        Sequence 15, Appl
       45      490    36.5      293    9   US-09-323-998D-43        Sequence 43, Appl


                                ALIGNMENTS


RESULT 1
US-10-320-797-3065
```

```
                        GenCore version 5.1.6
                  Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          April 15, 2004, 13:37:36 ; Search time 51.8 Seconds
                                           (without alignments)
                                           1412.738 Million cell updates/sec

Title:          US-08-737-319②  in PRT databases
Perfect score:  1398
Sequence:       1 MQLLAEDRTDHMRGASTWAG..........AALNTDKHEDWGTVHHINEA 259

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters:      1586107
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      A_Geneseq_29Jan04:*
                1:  geneseqp1980s:*
                2:  geneseqp1990s:*
                3:  geneseqp2000s:*
                4:  geneseqp2001s:*
                5:  geneseqp2002s:*
                6:  geneseqp2003as:*
                7:  geneseqp2003bs:*
                8:  geneseqp2004s:*
```

SEQ 2 in protein databases

```
        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.

                                SUMMARIES
                  %
Result          Query
  No.    Score  Match Length DB   ID                          Description
-------------------------------------------------------------------------------
    1    1398  100.0   259   2   AAW00279 WO 96/28545      Aaw00279 Peptide u
    2    1383   98.9   293   2   AAW35351 WO 97/36998      Aaw35351 Haematoco
    3    1383   98.9   293   3   AAY54300 WO 99/63055      Aay54300 Amino aci
    4    1383   98.9   293   3   AAY54288                  Aay54288 An isopen
    5    1372   98.1   293   5   ABG93893 USPAP 2002102631 Abg93893 H. pluvia
    6    1372   98.1   293   5   ABG93911                  Abg93911 Haematoco
    7    1368   97.9   305   2   AAW35350                  Aaw35350 Haematoco
    8    1368   97.9   305   3   AAY54287                  Aay54287 An isopen
    9    1360   97.3   305   5   ABG93892                  Abg93892 H. pluvia
   10  1352.5   96.7   304   3   AAY54301                  Aay54301 Amino aci
   11  1345.5   96.2   304   5   ABG93912                  Abg93912 Haematoco
   12   777.5   55.6   307   3   AAY54302                  Aay54302 Amino aci
   13   775.5   55.5   307   5   ABG93913                  Abg93913 Chlamydom
```

Appl.   earliest 03/1996   earliest 06/1998

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2004, 13:50:32 ; Search time 14.2196 Seconds
                                            (without alignments)
                                            1752.060 Million cell updates/sec

Title: US-08-737-319-2
Perfect score: 1398
Sequence: 1 MQLLAEDRTDHMRGASTWAG..........AALNTDKHEDWGTVHHINEA 259

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   PIR_78:*
                 1: pir1:*
                 2: pir2:*
                 3: pir3:*
                 4: pir4:*

        Pred. No. is the number of results predicted by chance to have a
        score greater than or equal to the score of the result being printed,
        and is derived by analysis of the total score distribution.

                              SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	1398	100.0	293	2	T46812	03/2000	isopentenyl-diphos
2	1394	99.7	293	2	T51248	07/2000	isopentenyl-diphos
3	1368	97.9	305	2	T52027	11/2000	isopentenyl-diphos
4	777.5	55.6	307	2	T07979		isopentenyl-diphos
5	530	37.9	290	2	S49588		isopentenyl-diphos
6	523	37.4	233	2	S71369		isopentenyl-diphos
7	522	37.3	227	2	A56442		isopentenyl-diphos
8	515	36.8	284	2	S71370		isopentenyl-diphos
9	496	35.5	288	2	A34440		isopentenyl-diphos
10	495.5	35.4	228	2	A53028		isopentenyl-diphos
11	398	28.5	831	2	S44843		K06H7.3 protein -
12	264	18.9	219	2	T03399		isopentenyl-diphos
13	255.5	18.3	197	2	T35275		probable IPP isome

```
                          GenCore version 5.1.6
                 Copyright (c) 1993 - 2004  Compugen Ltd.


OM protein - protein search, using sw model

Run on:          April 15, 2004, 13:46:31 ; Search time 10.1569 Seconds
                                             (without alignments)
                                             1327.789 Million cell updates/sec


Title:           US-08-737-319-2
Perfect score:   1398
Sequence:        1 MQLLAEDRTDHMRGASTWAG..........AALNTDKHEDWGTVHHINEA 259


Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5


Searched:        141681 seqs, 52070155 residues


Total number of hits satisfying chosen parameters:       141681


Minimum DB seq length: 0
Maximum DB seq length: 2000000000


Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries


Database :       SwissProt_42:*


     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                SUMMARIES
                  %
Result          Query
   No.   Score  Match Length DB  ID                      Description
     ------------------------------------------------------------------------
       1    533   38.1    235  1  IDI1_CAMAC              O48964 camptotheca
       2    530   37.9    287  1  IDI1_CLABR              Q39472 clarkia bre
       3  526.5   37.7    309  1  IDI2_CAMAC              O48965 camptotheca
       4    526   37.6    233  1  IDI1_ARATH              Q38929 arabidopsis
       5  524.5   37.5    286  1  IDI2_CLABR              Q39471 clarkia bre
       6    522   37.3    227  1  IDI1_SCHPO              Q10132 schizosacch
       7    515   36.8    284  1  IDI2_ARATH              Q42553 arabidopsis
       8    515   36.8    290  1  IDI2_CLAXA              Q39664 clarkia xan
       9    496   35.5    287  1  IDI1_YEAST              P15496 saccharomyc
      10  493.5   35.3    227  1  IDI1_HUMAN              Q13907 homo sapien
      11    491   35.1    251  1  IDI1_PHARH              O42641 phaffia rho
      12    473   33.8    227  1  IDI1_MOUSE              P58044 mus musculu
      13    472   33.8    227  1  IDI1_RAT                O35760 rattus norv
      14  471.5   33.7    227  1  IDI1_MESAU              O35586 mesocricetu
      15    415   29.7    227  1  IDI2_HUMAN              Q9bxs1 homo sapien
      16    266   19.0    197  1  IDI_STRAW               Q82mj7 streptomyce
      17  255.5   18.3    197  1  IDI_STRCO               Q9x7q6 streptomyce
```

```
                          GenCore version 5.1.6
               Copyright (c) 1993 - 2004  Compugen Ltd.


OM protein - protein search, using sw model

Run on:        April 15, 2004, 13:47:16 ; Search time 36.5647 Seconds
                                         (without alignments)
                                         2234.921 Million cell updates/sec


Title:         US-08-737-319-2
Perfect score: 1398
Sequence:      1 MQLLAEDRTDHMRGASTWAG.........AALNTDKHEDWGTVHHINEA 259


Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5


Searched:      1017041 seqs, 315518202 residues


Total number of hits satisfying chosen parameters:      1017041


Minimum DB seq length: 0
Maximum DB seq length: 2000000000


Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries


Database :     SPTREMBL_25:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
               10:  sp_plant:*
               11:  sp_rodent:*
               12:  sp_virus:*
               13:  sp_vertebrate:*
               14:  sp_unclassified:*
               15:  sp_rvirus:*
               16:  sp_bacteriap:*
               17:  sp_archeap:*


     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	1398	100.0	293	10	Q9SAY0	05/2000 or 10/1998	Q9say0 haematococc
2	1394	99.7	293	10	O81660	10/1998	O81660 haematococc
3	1368	97.9	305	10	O81659	10/1998	O81659 haematococc
4	777.5	55.6	307	10	O81691		O81691 chlamydomon
5	538	38.5	234	3	Q8NJL9		Q8njl9 aspergillus
6	534	38.2	268	3	Q874I0		Q874i0 emericella
7	530.5	37.9	226	10	Q8S3L7		Q8s3l7 melaleuca a
8	527	37.7	238	10	Q8GVZ0		Q8gvz0 oryza sativ
9	527	37.7	254	10	Q9M6K4		Q9m6k4 oryza sativ
10	527	37.7	263	10	Q9M6K6		Q9m6k6 lactuca sat
11	524	37.5	235	10	Q9AVG7		Q9avg7 nicotiana t
12	523	37.4	262	10	Q9FV48		Q9fv48 tagetes ere
13	522	37.3	245	10	Q9M6K5		Q9m6k5 tagetes ere
14	522	37.3	284	10	Q9M6K3		Q9m6k3 arabidopsis
15	521	37.3	261	10	Q9M6K2		Q9m6k2 arabidopsis
16	520	37.2	224	10	Q9FXR6		Q9fxr6 eucommia ul
17	519.5	37.2	235	10	Q8S3L8		Q8s3l8 melaleuca a
18	517	37.0	234	10	Q9S7C4		Q9s7c4 hevea brasi
19	517	37.0	267	10	Q9M6K8		Q9m6k8 adonis pala
20	517	37.0	290	10	Q9AVG8		Q9avg8 nicotiana t
21	516	36.9	280	10	Q9M6K7		Q9m6k7 lactuca sat
22	510	36.5	295	10	Q9M6K9		Q9m6k9 adonis pala
23	509	36.4	235	10	Q84RZ8		Q84rz8 periploca s
24	509	36.4	281	10	Q9M592		Q9m592 brassica ol
25	501	35.8	256	5	Q9VDC2		Q9vdc2 drosophila
26	500	35.8	256	5	Q8MRH1		Q8mrh1 drosophila
27	495.5	35.4	228	4	Q86U81		Q86u81 homo sapien
28	494	35.3	251	3	Q9UUQ1		Q9uuq1 phaffia rho
29	455.5	32.6	239	5	Q9NH02		Q9nh02 dictyosteli
30	398	28.5	236	5	Q9BI81		Q9bi81 caenorhabdi
31	379	27.1	227	11	Q8BFZ6		Q8bfz6 mus musculu
32	283	20.2	120	10	Q9AVI8		Q9avi8 youngia jap
33	276	19.7	120	10	Q9AYS6		Q9ays6 sonchus ole
34	275	19.7	120	10	Q94IE1		Q94ie1 euphorbia t
35	273	19.5	120	10	Q9AVJ5		Q9avj5 sonchus ole
36	272	19.5	229	5	Q8SSI5		Q8ssi5 encephalito
37	267	19.1	120	10	Q9AVG5		Q9avg5 sapium sebi
38	264	18.9	219	10	P93355		P93355 nicotiana t
39	256.5	18.3	121	10	Q9AVG6		Q9avg6 taraxacum j
40	243.5	17.4	119	10	Q9AVJ1		Q9avj1 metaplexis
41	236.5	16.9	197	2	Q8RTW6		Q8rtw6 uncultured
42	225.5	16.1	91	10	Q9LLB6		Q9llb6 daucus caro
43	225	16.1	192	2	Q8KZ66		Q8kz66 uncultured
44	204.5	14.6	503	16	Q83DT5		Q83dt5 coxiella bu
45	192	13.7	172	2	Q8KZ12		Q8kz12 uncultured

ALIGNMENTS

RESULT 1
Q9SAY0

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 15, 2004, 13:49:12 ; Search time 16.7588 Seconds
                                           (without alignments)
                                           797.856 Million cell updates/sec

Title: US-08-737-319-2
Perfect score: 1398
Sequence: 1 MQLLAEDRTDHMRGASTWAG..........AALNTDKHEDWGTVHHINEA 259

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_AA:*
                 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
                 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
                 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
                 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
                 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES   _Cunningham
                 %   family
Result   Query
   No.   Score Match Length DB   ID   Description
----
       1   1383   98.9   293   1   US-08-624-125-15 5744341   Sequence 15, Appl
       2   1383   98.9   293   4   US-08-937-155-15 652481)   Sequence 15, Appl
       3   1383   98.9   293   4   US-09-323-998E-15 664202) Sequence 15, Appl
       4   1383   98.9   293   4   US-09-323-998E-43   )"   Sequence 43, Appl
       5   1368   97.9   305   1   US-08-624-125-14   Sequence 14, Appl
       6   1368   97.9   305   4   US-08-937-155-14   Sequence 14, Appl
       7   1368   97.9   305   4   US-09-323-998E-14   Sequence 14, Appl
       8   1368   97.9   305   4   US-09-323-998E-44   Sequence 44, Appl
       9   777.5   55.6   307   4   US-09-323-998E-45   Sequence 45, Appl
      10   530   37.9   287   1   US-08-624-125-17   Sequence 17, Appl
      11   530   37.9   287   4   US-08-937-155-17   Sequence 17, Appl

```
                       GenCore version 5.1.6
                Copyright (c) 1993 - 2004  Compugen Ltd.


OM protein - protein search, using sw model

Run on:              April 15, 2004, 13:54:22 ; Search time 39.1039 Seconds
                                               (without alignments)
                                               1825.887 Million cell updates/sec

Title:               US-08-737-319-2
Perfect score:       1398
Sequence:            1 MQLLAEDRTDHMRGASTWAG.........AALNTDKHEDWGTVHHINEA 259

Scoring table:       BLOSUM62
                     Gapop 10.0 , Gapext 0.5

Searched:            1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters:       1124875

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :           Published_Applications_AA:*
                1:   /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
                2:   /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
                3:   /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
                4:   /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
                5:   /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
                6:   /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
                7:   /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
                8:   /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
                9:   /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
                10:  /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
                11:  /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
                12:  /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
                13:  /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
                14:  /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
                15:  /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
                16:  /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
                17:  /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
                18:  /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.
```

SUMMARIES

handwritten annotation, top right: USPAP 2004 0025202 earliest 6/2002

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1398	100.0	293	16	US-10-389-566-2125	Sequence 2125, Ap
2	1394	99.7	293	16	US-10-389-566-2102	Sequence 2102, Ap
3	1383	98.9	293	9	US-09-323-998D-15	Sequence 15, Appl
4	1383	98.9	293	9	US-09-323-998D-43	Sequence 43, Appl
5	1368	97.9	305	9	US-09-323-998D-14	Sequence 14, Appl
6	1368	97.9	305	9	US-09-323-998D-44	Sequence 44, Appl
7	1368	97.9	305	16	US-10-389-566-2101	Sequence 2101, Ap
8	777.5	55.6	307	9	US-09-323-998D-45	Sequence 45, Appl
9	538	38.5	271	14	US-10-128-714-3424	Sequence 3424, Ap
10	530.5	37.9	226	16	US-10-389-566-1492	Sequence 1492, Ap
11	530	37.9	287	9	US-09-323-998D-17	Sequence 17, Appl
12	527	37.7	229	9	US-09-323-998D-38	Sequence 38, Appl
13	527	37.7	238	9	US-09-323-998D-41	Sequence 41, Appl
14	527	37.7	238	16	US-10-389-566-1985	Sequence 1985, Ap
15	527	37.7	263	16	US-10-389-566-2297	Sequence 2297, Ap
16	526	37.6	271	14	US-10-128-714-8424	Sequence 8424, Ap
17	523	37.4	262	16	US-10-389-566-2458	Sequence 2458, Ap
18	522	37.3	245	16	US-10-389-566-2298	Sequence 2298, Ap
19	522	37.3	284	9	US-09-323-998D-16	Sequence 16, Appl
20	521	37.3	233	9	US-09-323-998D-42	Sequence 42, Appl
21	521	37.3	261	9	US-09-323-998D-18	Sequence 18, Appl
22	519.5	37.2	235	16	US-10-389-566-1491	Sequence 1491, Ap
23	518	37.1	232	9	US-09-323-998D-36	Sequence 36, Appl
24	517	37.0	234	9	US-09-323-998D-40	Sequence 40, Appl
25	517	37.0	234	16	US-10-389-566-2210	Sequence 2210, Ap
26	517	37.0	267	16	US-10-389-566-2295	Sequence 2295, Ap
27	516	36.9	280	9	US-09-323-998D-37	Sequence 37, Appl
28	516	36.9	280	16	US-10-389-566-2296	Sequence 2296, Ap
29	515	36.8	284	14	US-10-342-224-96	Sequence 96, Appl
30	515	36.8	290	16	US-10-389-566-2247	Sequence 2247, Ap
31	513	36.7	284	14	US-10-032-585-7474	Sequence 7474, Ap
32	510	36.5	295	16	US-10-389-566-2294	Sequence 2294, Ap
33	510	36.5	309	16	US-10-389-566-792	Sequence 792, App
34	509	36.4	281	16	US-10-389-566-2306	Sequence 2306, Ap
35	504	36.1	295	9	US-09-323-998D-39	Sequence 39, Appl
36	496	35.5	288	9	US-09-323-998D-19	Sequence 19, Appl
37	496	35.5	288	15	US-10-369-493-22482	Sequence 22482, A
38	495.5	35.4	228	12	US-10-363-616-265	Sequence 265, App
39	488	34.9	281	15	US-10-320-797-3065	Sequence 3065, Ap
40	487	34.8	267	15	US-10-369-493-2346	Sequence 2346, Ap
41	487	34.8	267	15	US-10-369-493-2351	Sequence 2351, Ap
42	477	34.1	273	15	US-10-369-493-12865	Sequence 12865, A
43	476.5	34.1	257	15	US-10-369-493-3496	Sequence 3496, Ap
44	448	32.0	259	16	US-10-389-566-925	Sequence 925, App
45	445.5	31.9	194	9	US-09-925-302-585	Sequence 585, App

handwritten annotation, right of rows 3–6: issued

ALIGNMENTS

RESULT 1
US-10-389-566-2125